

НОВЫЙ МЕТОД ПРОГНОЗИРОВАНИЯ ВЕРОЯТНОСТИ ВЫЖИВАНИЯ И ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ ДЛЯ ПАЦИЕНТОВ С ТРАВМАМИ

В.Г. Щетинин^{1,2}, Л.И. Якайте², В.Ф. Куракин¹, В.И. Горбаченко¹

¹ ФГБОУ ВО «Пензенский государственный университет», Пенза, Российская Федерация

² Университет Бедфордшира, Лутон, Великобритания

Контактная информация: Виталий Георгиевич Щетинин, кандидат технических наук, старший преподаватель кафедры вычислительной техники Университета Бедфордшира.

E-mail: vitaly.schetinin@beds.ac.uk

АННОТАЦИЯ

Для оценки тяжести травм и повреждений в настоящее время используется «золотой» стандарт *Trauma and Injury Severity Score (TRISS)*, предназначенный для скрининга состояния пациента с целью предсказания вероятности выживания. Однако использование этого стандарта в течение более чем 40 лет выявило ряд проблем: первое — необъяснимую флуктуацию предсказанных значений, вызванную агрегированием скрининговых тестов, второе — недостаточную точность оценки интервалов неопределенности, в которых распределены предсказания.

Для снижения негативного влияния этих факторов нами разработан новый метод и сделан доступным для практиков в виде *web*-калькулятора. Метод использует байесовскую методологию статистического вывода, которая теоретически позволяет достичь максимальной точности предсказаний, являясь, однако, вычислительно сложной в реализации. Метод был реализован и верифицирован нами на выборке данных, включающей 571 148 пациентов, зарегистрированных в *US National Trauma Data Bank (NTDB)*, с числом травм от 1 до 20. Распределение пациентов по группам по числу травм: 1-я группа — 174 647 пациентов имели 1 травму, 2-я группа — 381 137 — от 2 до 10 травм и 3-я группа 15 364 — от 11 до 20 травм. Доли выживших в каждой категории были 0,977, 0,953 и 0,831 соответственно. Предложенный нами метод улучшил точность предсказаний на 0,04%, 0,36% и 3,64% (значимость $p < 0,05$) для каждой указанной группы. Критерий Хосмер-Лемешоу показал значительное улучшение калибровки новой модели. Интервалы неопределенности 2σ были снижены с 0,628 до 0,569 для пациентов 2-й группы и с 1,227 до 0,930 для пациентов 3-й группы ($p < 0,005$). Новый метод показал статистически значимое улучшение как точности предсказания выживания, так и точности оценки интервалов неопределенности. Наибольшее улучшение достигается для пациентов 3-й группы. Метод сделан доступным для практиков как *web*-калькулятор <http://www.traumacalc.org>.

Ключевые слова:

оценка тяжести травм и повреждений, *TRISS*, предсказание выживания, *web*-калькулятор.

TRISS — *Trauma and Injury Severity Score*
NTDB — *National Trauma Data Bank*

ISS — *Injury Severity Score*
RTS — *Revised Trauma Score*

ВВЕДЕНИЕ

Вероятность выживания пациента, доставленного в стационар, оценивается по модели *TRISS (Trauma and Injury Severity Score)* [1–3]. Эта модель является моделью логистической регрессии, в которой используется до 3 наиболее тяжелых повреждений в 6 областях тела: голова, лицо, грудная клетка, брюшная полость, конечности и поверхность (кожа, подкожная клетчатка и ожоги). Модель включает как непрерывные, так и категориальные скрининговые тесты (или предикторы). Непрерывные предикторы включают такие показатели, как возраст, систолическое артериальное давление, частота дыхания. Категориальные предикторы включают: оценку тяжести травмы, полученной пациентом, шкалу *GCS (Glasgow Coma Scale, Шкала комы Глазго)*, а также тип повреждения (тупое или проникающее). Все эти параметры оцениваются обученным специалистом сразу по прибытии пациента. Вычисление вероятностей выполняется с помощью метода, реализованного в виде доступного онлайн *TRISS Calculator* [4].

Скрининговые тесты образуют два агрегированных предиктора: показатель тяжести травм (*ISS, Injury Severity Score*) и модифицированная шкала оценки травмы (*RTS, Revised Trauma Score*). Практики обнаружили, что при такой агрегации возникают необъяснимые флуктуации *ISS* в предсказываемых вероятностях выживания [5], что влияет на точность прогнозов.

Формально модель *TRISS* определяет вероятность выживания, P_s , используя следующую логистическую регрессию: $P_s = 1/(1 + e^{-b})$, где $b = b_0 + b_1 \times RTS + b_2 \times ISS + b_3 \times A$. Здесь b_0 , b_1 , b_2 , и b_3 — параметры регрессии, а $A = 0$, если возраст пациента менее 55 лет, и $A = 1$ в противном случае. Параметры регрессии были определены для тупых и проникающих ранений.

Анализ статистики выживаемости и предсказанных вероятностей определяется как калибровка, а визуализация отношений между этими вероятностями является калибрационной кривой [6]. С этой точки зрения калибрационная кривая для метода *TRISS* имеет

значительное отклонение от фактических вероятностей выживания [7].

Точность предсказаний, базирующихся на вычислениях, полученных с помощью шкалы *TRISS*, признается удовлетворительной, когда типы и тяжесть травм типичны. Однако для пациентов с числом травм более трех, а также с некоторыми нетипичными комбинациями травм, точность недостаточна, и могла бы быть улучшена [3].

На практике крайне важно как можно более точно оценить неопределенность в прогнозах выживания. Оценки неопределенности необходимы для того, чтобы минимизировать риски фатальных ошибок. В частности, неопределенность может быть представлена в виде доверительных интервалов, которые могут быть надежно оценены, когда известна прогнозируемая плотность вероятности, что достижимо в ограниченном числе тривиальных ситуаций. Оценки распределения в практических случаях невозможны в рамках методологии *TRISS*, так как эта методология основана на теоретическом предположении о распределении данных, требуемом для логистической регрессии [8].

Для уменьшения негативного влияния вышеуказанных проблем нами разработан новый метод прогнозирования выживаемости. Наш подход основан на байесовской методологии статистического вывода, который в теории обеспечивает наиболее точные прогнозы. До недавнего времени его использование в практике было ограничено ввиду вычислительной сложности. Байесовские методы обеспечивают практиков наиболее точными предсказаниями и оценками интервалов неопределенности, в пределах которых распределяются прогнозируемые значения [9–11]. Нами разработан удобный для применения в любых условиях *web*-калькулятор для прогнозирования выживаемости [12]. Основные результаты этого исследования были опубликованы в 2013 г. [13].

МАТЕРИАЛ И МЕТОДЫ

Предложенный метод основан на Байесовском подходе к построению вероятностных зависимостей выживания пациентов от их физиологических состояний, представленных значениями предикторов (или независимых переменных). Искомая зависимость представляется в виде модели, включающей параметры, которые определяются на имеющейся выборке данных. Байесовский подход позволяет найти исковую модель путем усреднения некоторого множества моделей, которые могут быть созданы для объяснения заданных данных [11, 12].

Модели оцениваются качеством предсказания на заданной выборке данных в процессе калибровки. В медицинской практике качество моделей обычно оценивается критерием Хосмер–Лемешоу (*HL*) [6, 7]. Чем меньше значение *HL*-критерия, тем лучше калибровка и, следовательно, будет выше точность прогноза. Мы использовали *HL*-критерий для сравнения предлагаемого подхода и стандартного *TRISS* на выборке данных пациентов, зарегистрированных в США в Национальном банке данных о травме (*NTDB*), который является основным источником данных записей травмированных пациентов, поступивших в лечебные учреждения [2]. Данные *NTDB* включают в себя информацию о пациенте, такую как возраст, пол, типы травмы, а также некоторые клинические данные о состоянии пациента. В *NTDB* также включена информация о

прогнозе *TRISS* и результатах лечения (жив или умер) для каждого пациента.

Для тестирования нашего подхода мы использовали данные о 571 148 пациентах, зарегистрированных в *NTDB* с 1–20 травмами. Эти пациенты были распределены на 3 категории телесных повреждений следующим образом: 1-я группа — 174 647 пациентов с одной травмой, 2-я группа — 381 137 — с 2–10, 3-я группа — 15 364 с 11–20 травмами. Показатели выживаемости в каждой категории были 0,977; 0,953 и 0,831 соответственно.

Таблица

Скрининг-тесты (*Screening tests*) и диапазоны значений (*Range*)

N	Screening tests (Скрининг-тесты)	Range (Диапазон значений)
1	Age (Возраст)	0–100
2	Gender (Пол)	0 female (женский), 1 male (мужской)
3	Injury type (Тип повреждения)	0 penetrating (проникающее), 1 blunt (тупое), 2 (ожог)
4	Blood pressure (Артериальное давление)	0–270
5	Respiration rate (Частота дыхания)	0–100
6	GCS Eye (Открытие глаз)	0–5
7	GCS Verbal (Речевая реакция)	0–5
8	GCS Motor (Двигательная реакция)	0–6
9	Head severity (Тяжесть повреждения, голова)	0–6
10	Face severity (Тяжесть повреждения, лицо)	0–6
11	Neck severity (Тяжесть повреждения, шея)	0–6
12	Thorax severity (Тяжесть повреждения, грудная клетка)	0–9
13	Spine severity (Тяжесть повреждения, позвоночник)	0–9
14	Abdomen severity (Тяжесть повреждения, брюшная полость)	0–6
15	Upper extremity severity (Тяжесть повреждения, верхние конечности)	0–6
16	Lower extremity severity (Тяжесть повреждения, нижние конечности)	0–6
17	External severity (Тяжесть повреждения, поверхность тела)	0–6
18	Discharge status (Статус)	1 alive (жив), 0 dead (умер)

Таблица показывает скрининг-тесты (или предикторы), которые используются в методе *TRISS*. Типы переменных — возраст (*Age*), артериальное давление (*Blood pressure*) и частота дыхания (*Respiration rate*) являются непрерывными, остальные предикторы являются категориальными. Предсказанный исход — это статус пациента 1 — жив (*alive*), или 0 — умер (*dead*). Таблица также показывает диапазоны значений каждого предиктора.

РЕЗУЛЬТАТЫ

Калибровочные кривые были рассчитаны для 10 интервалов, равномерно распределенных по вероятности выживания от 0,0 до 1,0. На рис. 1 калибровочные кривые представлены для стандартного и предложенного методов. Доверительные интервалы показаны черными линиями, а наблюдаемое (фактическое) выживание в каждом интервале вероятностей выживания обозначается красной звездочкой. Как видно, эти кривые существенно отличаются по точности оценки доверительных интервалов. Калибровочная

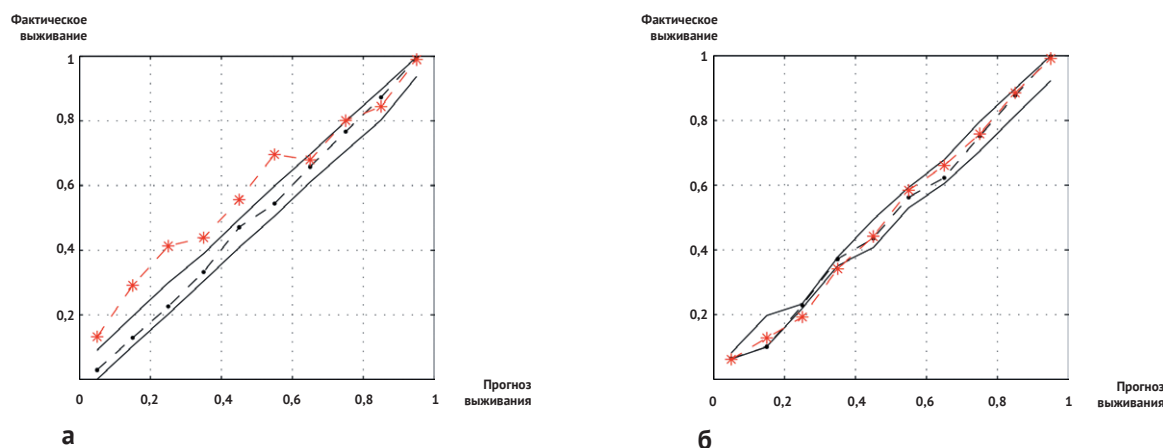


Рис. 1. Калибровочные кривые и доверительные интервалы для TRISS (а) и предложенного метода (б)

кривая для TRISS показывает, что большинство наблюдаемых вероятностей выживания лежит вне доверительного интервала. Наоборот, на калибровочной кривой, представленной на правом графике рис. 1, все наблюдаемые вероятности выживания лежат в пределах доверительного интервала, предсказанного предложенным методом. Это свидетельствует о высокой точности оценки интервалов. Значения *HL*-критерия значительно отличаются и равны 3893,4 и 557,1 для TRISS и нового метода соответственно.

Предложенный и стандартный (TRISS) методы были проанализированы с точки зрения точности классификации. Прогнозы с вероятностями выживания более 0,5 относились к классу «выживших», а результаты прогнозов с меньшими вероятностями относились к классу «умерших».

Предсказания, сделанные обоими методами для пациентов с числом повреждений от 1 до 20, представлены на рис. 2. Можно видеть, что прогнозы по методу TRISS (синие звездочки) значительно хуже предсказаний, сделанных предложенным методом (красные треугольники). Различия между предсказаниями TRISS и фактическими вероятностями постепенно возрастают с ростом числа повреждений, в то время как различие между байесовским прогнозом и фактическими вероятностями остается небольшим.

В целом новый метод улучшает точность предсказания на 0,04%, 0,36% и 3,64% ($p < 0,05$) для пациентов с 1, 2–10 и 11–20 повреждениями соответственно. *HL*-критерий показал значительное улучшение согласия новой модели с реальными данными. В терминах неопределенности прогнозирования 2σ интервалы были сокращены с 0,628 до 0,569 для пациентов с 2–10 травмами и с 1,227 до 0,930 для пациентов с 11–20 повреждениями, причём оба результата оказались статистически значимыми ($p < 0,005$).

ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

Таким образом, «золотой» стандарт TRISS, используемый для оценки тяжести травмы в течение более 40 лет, был улучшен нами с точки зрения точности

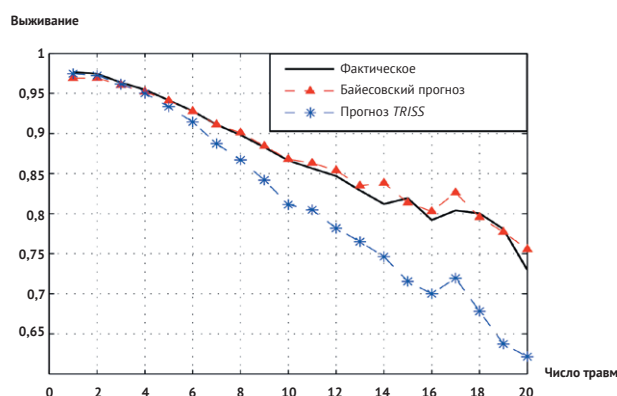


Рис. 2. Сравнение прогнозов, сделанных методом TRISS и предложенным методом (Bayesian) с фактическими вероятностями при разном числе травм (Number of injuries)

прогноза выживаемости и оценки интервалов неопределенности. Методологически TRISS не способен дать оценки распределения вероятностей выживания, которые востребованы для определения доверительных интервалов индивидуально для каждого пациента. За время использования обнаружено, что агрегация скрининговых тестов, выполняемая по шкале TRISS, приводит к необъяснимым колебаниям и ошибочным решениям. Новый подход, разработанный для прогноза выживания, был верифицирован на данных о пациентах, зарегистрированных в NTDB, и показал статистически значимое улучшение точности прогнозирования и оценок интервалов неопределенности. Наибольшее улучшение было достигнуто у больных с 11–20 травмами. Новый метод доступен для лечащих врачей в виде онлайн-калькулятора TraumaCalc, <http://www.traumacalc.org/>.

БЛАГОДАРНОСТИ

Авторы благодарны Prof Lecky и Dr Bouamra из UK Trauma Audit and Research Network, University of Manchester, за вдохновляющие обсуждение.

ЛИТЕРАТУРА

- Boyd C.R., Tolson M.A., Copes W.S. Evaluating trauma care: The TRISS method. Trauma Score and the Injury Severity Score. J. Trauma. 1987; 27(4): 370–378. PMID: 310664.
- Committee on Trauma, American College of Surgeons NTDB Research Data Set and National Sample Program. 2014. URL: <http://www.facs.org/trauma/ntdb/ntdbapp.html>. Accessed Jul 01, 2016.
- Kilgo P., Meredith J., Osler T. Injury severity scoring and outcomes research. In: Trauma / eds. D.V. Feliciano, K.L. Mattox, E.E. Moore. 6th ed. New York: McGraw-Hill, 2008. 223–230.
- Brohi K. TRISS: Trauma — injury severity score. TRISS — Overview and Desktop Calculator. 2012. URL: <http://www.trauma.org/index.php/main/article/387>. Accessed Jul 01, 2016.

5. Osler T., Glance L., Buzas J., et al. A trauma mortality prediction model based on the anatomic injury scale. *Ann. Surg.* 2008; 247(6): 1041–1048. DOI: 10.1097/SLA.0b013e31816ffb3f.
6. Steyerberg E., Vickers A., Cook N., et al. Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology.* 2010; 21(1): 128–138. DOI: 10.1097/EDE.0b013e3181c30fb2.
7. Rogers F., Osler T., Krasne M., et al. Has TRISS become an anachronism? A comparison of mortality between the National Trauma Data Bank and Major Trauma Outcome Study databases. *J. Trauma Acute Care Surg.* 2012; 73(2): 326–331. DOI: 10.1097/TA.0b013e31825a7758.
8. Bailey T.C., Everson R.M., Fieldsend J.E., et al. Representing classifier confidence in the safety critical domain — an illustration from mortality prediction in trauma cases. *Neur. Comput. Applicat.* 2007; 16(1): 1–10. DOI: 10.1007/s00521-006-0053-y.
9. Negrin M.A., Nam J., Briggs A.H. Bayesian solutions for handling uncertainty in survival extrapolation. *Med. Decis. Making.* 2016. pii: 0272989X16650669. DOI: 10.1177/0272989X16650669.
10. Schetinin V., Jakaite L., Krzanowski W.J. Bayesian prediction for survival of patients: Study on the US national trauma data bank. *Comput. Methods Programs Biomed.* 2013; 111(3): 602–612. DOI: 10.1016/j.cmpb.2013.05.015.
11. Schetinin V., Jakaite L., Krzanowski W.J. Prediction of survival probabilities with Bayesian decision trees. *Exp. Syst. Applicat.* 2013; 40(14): 5466–5476. DOI: 10.1016/j.eswa.2013.04.009.
12. Schetinin V., Jakaite L. TraumaCalc. Bayesian prediction of trauma survival. 2016. URL: <http://www.traumacalc.org> Accessed Jul 01, 2016.
13. Schetinin V., Jakaite L., Jakaitis J., Krzanowski W. Bayesian decision trees for predicting survival of patients: A study on the US national trauma data bank. *Computer Methods and Programs in Biomedicine.* 2013; 111(3): 602–612. DOI: <http://dx.doi.org/10.1016/j.cmpb.2013.05.015>.

REFERENCES

1. Boyd C.R., Tolson M.A., Copes W.S. Evaluating trauma care: The TRISS method. Trauma Score and the Injury Severity Score. *J. Trauma.* 1987; 27(4): 370–378. PMID: 310664.
2. Committee on Trauma, American College of Surgeons. NTDB Research Data Set and National Sample Program. 2014. URL: <http://www.facs.org/trauma/ntdb/ntdbapp.html>. (Accessed Jul 01, 2016.)
3. Kilgo P., Meredith J., Osler T. Injury severity scoring and outcomes research. In: Feliciano D.V., Mattox K.L., Moore E.E., eds. *Trauma*. 6th ed. New York: McGraw-Hill, 2008: 223–230.
4. Brohi K. TRISS: Trauma — injury severity score. TRISS — Overview and Desktop Calculator. 2012. URL: <http://www.trauma.org/index.php/main/article/387>. (Accessed Jul 01, 2016.)
5. Osler T., Glance L., Buzas J., et al. A trauma mortality prediction model based on the anatomic injury scale. *Ann. Surg.* 2008; 247(6): 1041–1048. DOI: 10.1097/SLA.0b013e31816ffb3f.
6. Steyerberg E., Vickers A., Cook N., et al. Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology.* 2010; 21(1): 128–138. DOI: 10.1097/EDE.0b013e3181c30fb2.
7. Rogers F., Osler T., Krasne M., et al. Has TRISS become an anachronism? A comparison of mortality between the National Trauma Data Bank and Major Trauma Outcome Study databases. *J. Trauma Acute Care Surg.* 2012; 73(2): 326–331. DOI: 10.1097/TA.0b013e31825a7758.
8. Bailey T.C., Everson R.M., Fieldsend J.E., et al. Representing classifier confidence in the safety critical domain — an illustration from mortality prediction in trauma cases. *Neur. Comput. Applicat.* 2007; 16(1): 1–10. DOI: 10.1007/s00521-006-0053-y.
9. Negrin M.A., Nam J., Briggs A.H. Bayesian solutions for handling uncertainty in survival extrapolation. *Med. Decis. Making.* 2016. PII: 0272989X16650669. DOI: 10.1177/0272989X16650669.
10. Schetinin V., Jakaite L., Krzanowski W.J. Bayesian prediction for survival of patients: Study on the US national trauma data bank. *Comput. Methods Programs Biomed.* 2013; 111(3): 602–612. DOI: 10.1016/j.cmpb.2013.05.015.
11. Schetinin V., Jakaite L., Krzanowski W.J. Prediction of survival probabilities with Bayesian decision trees. *Exp. Syst. Applicat.* 2013; 40(14): 5466–5476. DOI: 10.1016/j.eswa.2013.04.009.
12. Schetinin V., Jakaite L. TraumaCalc. Bayesian prediction of trauma survival. 2016. URL: <http://www.traumacalc.org> (Accessed Jul 01, 2016.)
13. Schetinin V., Jakaite L., Jakaitis J., Krzanowski W. Bayesian decision trees for predicting survival of patients: A study on the US national trauma data bank. *Computer Methods and Programs in Biomedicine.* 2013; 111(3): 602–612. DOI: <http://dx.doi.org/10.1016/j.cmpb.2013.05.015>.

Конфликт интересов отсутствует.

Поступила 20.07.2016

A NEW METHOD FOR PREDICTING SURVIVAL AND ESTIMATING UNCERTAINTY IN TRAUMA PATIENTS

V.G. Schetinin^{1,2}, L.I. Jakaite², V.F. Kuriakin¹, V.I. Gorbachenko¹¹ Penza State University, Penza, Russian Federation² University of Bedfordshire, Luton, UKContacts: Vitaly Schetinin, PhD, Senior Lecturer in Computing and Information Systems, University of Bedfordshire. E-mail: vitaly.schetinin@beds.ac.uk

ABSTRACT The Trauma and Injury Severity Score (TRISS) is the current “gold” standard of screening patient’s condition for purposes of predicting survival probability. More than 40 years of TRISS practice revealed a number of problems, particularly, 1) unexplained fluctuation of predicted values caused by aggregation of screening tests, and 2) low accuracy of uncertainty intervals estimations. We developed a new method made it available for practitioners as a web calculator to reduce negative effect of factors given above. The method involves Bayesian methodology of statistical inference which, being computationally expensive, in theory provides most accurate predictions. We implemented and tested this approach on a data set including 571,148 patients registered in the US National Trauma Data Bank (NTDB) with 1–20 injuries. These patients were distributed over the following categories: (1) 174,647 with 1 injury, (2) 381,137 with 2–10 injuries, and (3) 15,364 with 11–20 injuries. Survival rates in each category were 0.977, 0.953, and 0.831, respectively. The proposed method has improved prediction accuracy by 0.04%, 0.36%, and 3.64% (p-value <0.05) in the categories 1, 2, and 3, respectively. Hosmer-Lemeshow statistics showed a significant improvement of the new model calibration. The uncertainty 2σ intervals were reduced from 0.628 to 0.569 for patients of the second category and from 1.227 to 0.930 for patients of the third category, both with p-value <0.005. The new method shows the statistically significant improvement (p-value <0.05) in accuracy of predicting survival and estimating the uncertainty intervals. The largest improvement has been achieved for patients with 11–20 injuries. The method is available for practitioners as a web calculator <http://www.traumacalc.org>.

Keywords: trauma care, trauma injury severity score, survival prediction.

DOI: 10.23934/2223-9022-2017-6-1-30-33